IDENTIFIABILITY OF HIDDEN MARKOV MODELS

BERLIAN SETIAWATY

Department of Mathematics,
Faculty of Mathematics and Natural Sciences,
Bogor Agricultural University
Jl. Meranti, Kampus IPB Darmaga, Bogor, 16680 Indonesia

Abstract. This paper shows that the identifiability problem for hidden Markov models can be derived from the identifiability of finite mixtures.

Keywords: Hidden Markov model, finite mixture, identifiable.

1. Introduction

The purpose of this paper is to show that, using slightly modification, the identifiability of hidden Markov models can be derived from the identifiability of finite mixtures which is already established (see [5]).

We will begin with definition of a hidden Markov model and its true parameter, then go to the identifiability problem. We will also present identifiability of finite mixtures and in the last section we show that the identifiability of hidden Markov models can be derived from the identifiability of finite mixtures.

2. A Hidden Markov Model and Its True Parameter

Precisely, according to [2], a hidden Markov model is formally defined as follows.

Definition 2.1. A pair of discrete time stochastic processes $\{(X_t, Y_t) : t \in \boldsymbol{N}\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ and taking values in a set $\boldsymbol{S} \times \mathcal{Y}$, is said to be a ***hidden Markov model*** (HMM), if it satisfies the following conditions.

1. $\{X_t\}$ is a finite state Markov chain.
2. Given $\{X_t\}$, $\{Y_t\}$ is a sequence of conditionally independent random variables.
3. The conditional distribution of Y_n depends on $\{X_t\}$ only through X_n.
4. The conditional distribution of Y_t given X_t does not depend on t.

Assume that the Markov chain $\{X_t\}$ ***is not observable***. The cardinality K of $\boldsymbol{S}$, will be called the ***size*** of the hidden Markov model.

From [3], it can be seen that the law of the hidden Markov model $\{(X_t, Y_t)\}$ is completely specified by :
(a). The size K.
(b). The transition probability matrix $A = (\alpha_{ij})$, satisfying

$$\alpha_{ij} \geq 0, \qquad \sum_{j=1}^{K} \alpha_{ij} = 1, \qquad i, j = 1, \dots, K.$$

(c). The initial probability distribution $\pi = (\pi_i)$ satisfying

$$\pi_i \geq 0, \qquad i = 1, \dots, K, \qquad \sum_{i=1}^{K} \pi_i = 1.$$

(d). The vector $\theta = (\theta_i)^T$, $\theta_i \in \Theta$, $i = 1, \dots, K$, which desribes the conditional
densities of Y_t given $X_t = i$, $i = 1, \dots, K$.

Definition 2.2. Let

$$\phi = (K, A, \pi, \theta).$$

The parameter ϕ is called a ***representation*** of the hidden Markov model
$\{(X_t, Y_t)\}$.

Thus, the hidden Markov model $\{(X_t, Y_t)\}$ can be represented by a representation $\phi = (K, A, \pi, \theta)$.

Let $\phi = (K, A, \pi, \theta)$ and $\widehat{\phi} = (\widehat{K}, \widehat{A}, \widehat{\pi}, \widehat{\theta})$ be two representations which respectively generate hidden Markov models $\{(X_t, Y_t)\}$ and $\{(\widehat{X}_t, Y_t)\}$. The $\{(X_t, Y_t)\}$ takes values on $\{1, \dots, K\} \times \mathcal{Y}$ and $\{(\widehat{X}_t, Y_t)\}$ takes values on $\{1, \dots, \widehat{K}\} \times \mathcal{Y}$. For any $n \in \boldsymbol{N}$, let $p_\phi(\cdot, \cdots, \cdot)$ and $p_{\widehat{\phi}}(\cdot, \cdots, \cdot)$ be the n-dimensional joint density function of $Y_1, \dots Y_n$ with respect to ϕ and $\widehat{\phi}$. Suppose that for every $n \in \boldsymbol{N}$,

$$p_\phi(Y_1, \dots, Y_n) = p_{\widehat{\phi}}(Y_1, \dots, Y_n).$$

Then $\{Y_t\}$ has the same law under ϕ and $\widehat{\phi}$. Since in hidden Markov models $\{(X_t, Y_t)\}$ and $\{(\widehat{X}_t, Y_t)\}$, the Markov chains $\{X_t\}$ and $\{\widehat{X}_t\}$ are not observable and we only observed the values of $\{Y_t\}$, then theoretically, the hidden Markov models $\{(X_t, Y_t)\}$ and $\{(\widehat{X}_t, Y_t)\}$ are *indistinguishable*. In this case, it is said that $\{(X_t, Y_t)\}$ and $\{(\widehat{X}_t, Y_t)\}$ are *equivalent*. The representations ϕ and $\widehat{\phi}$ are also said to be *equivalent*, and will be denoted as $\phi \sim \widehat{\phi}$.

For each $K \in \boldsymbol{N}$, define

$$\begin{aligned}\Phi_K \;=\; \Big\{ \phi : \phi = (K, A, \pi, \theta),\ &\text{where } A,\ \pi \text{ and } \theta \text{ satisfy :} \\ &A = (\alpha_{ij}), \quad \alpha_{ij} \geq 0, \quad \sum_{j=1}^{K} \alpha_{ij} = 1, \quad i, j = 1, \dots, K \\ &\pi = (\pi_i), \quad \pi_i \geq 0, \quad i = 1, \dots, K, \quad \sum_{i=1}^{K} \pi_i = 1 \\ &\theta = (\theta_i)^T, \quad \theta_i \in \Theta, \quad i = 1, \dots, K \Big\} \end{aligned} \tag{2.1}$$

and

$$\Phi = \bigcup_{K \in \boldsymbol{N}} \phi_K. \tag{2.2}$$

The relation $\sim$ is now defined on Φ as follows.

Definition 2.3. Let $\phi, \widehat{\phi} \in \Phi$. Representations ϕ and $\widehat{\phi}$ are said to be ***equivalent***, denoted as

$$\phi \sim \widehat{\phi}$$

if and only if for every $n \in \boldsymbol{N}$,

$$p_\phi(Y_1, Y_2, \dots, Y_n) = p_{\widehat{\phi}}(Y_1, Y_2, \dots, Y_n).$$

Remarks 2.4. It is clear that relation $\sim$ forms an equivalence relation on Φ.

Definition 2.5. Let $\{(X_t, Y_t)\}$ be a hidden Markov model with representation $\phi \in \Phi$. A representation $\phi^o = (K^o, A^o, \pi^o, \theta^o) \in \Phi$, is called ***a true parameter*** of the hidden Markov model $\{(X_t, Y_t)\}$ if and only if

1. $\phi^o \sim \phi$.
2. K^o is ***minimum***, that is, there is no $\widehat{\phi} \in \Phi_K$, with $K < K^o$, such that $\widehat{\phi} \sim \phi^o$.

3. Identifiability Problem

Let $\phi^o = (K^o, A^o, \pi^o, \theta^o)$ be a true parameter of a hidden Markov model $\{(X_t, Y_t)\}$. According to [4], if $\phi \in \Phi_K$ and $\phi \sim \phi^o$, then $K \geq K^o$. Moreover, there are infinitely many $\phi \in \Phi_K$, with $K > K_o$ and at least finitely many $\phi \in \Phi_K$, with $K = K^o$, such that $\phi \sim \phi^o$.

Let

$$\mathcal{T} = \left\{ \phi \in \cup_{K \geq K^o} \Phi_K : \phi \sim \phi^o \right\}.$$

For parameter estimation purposes, every $\phi \in \mathcal{T}$ must be *identifiable*. This means that all parameters of ϕ can be identified with parameters of ϕ^o.

Let $\phi = (K^o, A, \pi, \theta) \in \mathcal{T}$. Since $\phi \sim \phi^o$, then by definition for any $n \in \boldsymbol{N}$, the n-dimensional joint density functions of $Y_1, \ldots, Y_n$ under ϕ and ϕ^o are the same, that is,

$$p_{\phi^o}(y_1, \ldots, y_n) = p_\phi(y_1, \ldots, y_n), \tag{3.1}$$

for every $(y_1, \ldots, y_n) \in \mathcal{Y}^n$. Consider a special case of (3.1), when $n = 1$,

$$\begin{aligned} p_{\phi^o}(y_1) &= p_\phi(y_1) \\ \sum_{i=1}^{K^o} \pi_i^o f(y_1, \theta_i^o) &= \sum_{i=1}^{K^o} \pi_i f(y_1, \theta_i). \end{aligned} \tag{3.2}$$

From (3.2), we must be able to identify each (π_i, θ_i), with (π_j^o, θ_j^o). In other words, we must be able to show that for every i, $i = 1, \ldots, K^o$, there is j, $1 \le j \le K^o$, such that

$$\pi_i = \pi_j^o \qquad \text{and} \qquad \theta_i = \theta_j^o,$$

which can be written in the implication form,

$$\sum_{i=1}^{K^o} \pi_i f(y_1, \theta_i) = \sum_{i=1}^{K^o} \pi_i^o f(y_1, \theta_i^o) \quad \Longrightarrow \quad \begin{array}{l} \forall i,\ 1 \le i \le K^o,\ \ \exists j,\ 1 \le j \le K^o \\ \text{such that } \pi_i = \pi_i^o \text{ and } \theta_i = \theta_j^o. \end{array} \tag{3.3}$$

Consider the following example.

Example 3.1. Suppose that from the observation Y_1 has a density function as in Figure 1. Since we only observe the values of $\{Y_t\}$, then there is no way we can tell if the observation comes from

$$p(y_1) = \tfrac{1}{4} U(-1, 1) + \tfrac{3}{4} U(-3, 3)$$

or

$$p(y_1) = \tfrac{1}{2} U(-3, 1) + \tfrac{1}{2} U(-1, 3),$$

where $U(a, b)$ is a uniform distribution with range (a, b).

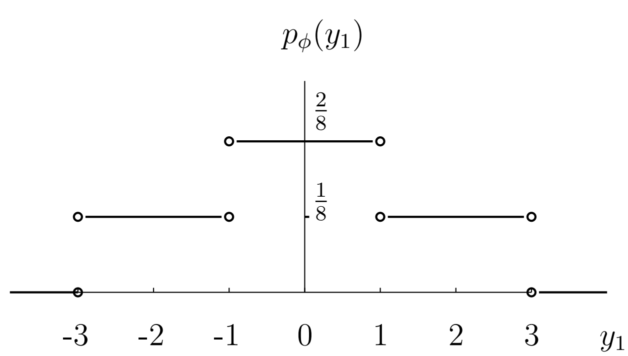

Figure 1.

The Example 3.1 above, shows that not every family of densities satisfies (3.3). Therefore we have to find conditions on the family of densities $\boldsymbol{\mathcal{F}} = \{f(\cdot,\theta) : \theta \in \Theta\}$, so that (3.3) holds.

Later, it can be shown, using a slight modification, we can apply identifiability criteria for finite mixtures, which have already been established (see [5]), to hidden Markov models, so it can be used to identify the true parameter ϕ^o.

4. Identifiability of Finite Mixtures

A formal definition of mixture distribution cited from [6] is as follows.

Definition 4.1. Let $\boldsymbol{\mathcal{F}} = \{F(\cdot,\theta) : \theta \in \mathcal{B}\}$ constitute a family of one dimensional distribution functions taking values in $\mathcal{Y}$ indexed by a point θ in a Borel subset $\mathcal{B}$ of Euclidean m-space $\boldsymbol{R}^m$, such that $F(\cdot,\cdot)$ is measurable in $\mathcal{Y} \times \mathcal{B}$. Let G be any distribution function such that the measure μ_G induced by G assigns measure 1 to $\mathcal{B}$. H is called a ***finite mixture*** if its mixing distribution G or rather the corresponding measure μ_G is discrete and assigns positive mass to only a finite number of points in $\mathcal{B}$. Thus the class $\widetilde{\boldsymbol{\mathcal{H}}}$ of finite mixtures on $\boldsymbol{\mathcal{F}}$ is defined by

$$\widetilde{\boldsymbol{\mathcal{H}}} = \left\{ H(\cdot) : H(\cdot) = \sum_{i=1}^{N} c_i F(\cdot,\theta_i),\ c_i > 0, \sum_{i=1}^{N} c_i = 1,\ F(\cdot,\theta_i) \in \boldsymbol{\mathcal{F}}, N = 1, 2 \ldots \right\}$$

that is, $\widetilde{\boldsymbol{\mathcal{H}}}$ is the convex hull of $\boldsymbol{\mathcal{F}}$.

Remarks 4.2. In every expression of finite mixture

$$H(\cdot) = \sum_{i=1}^{N} c_i F(\cdot,\theta_i),$$

$\theta_1, \ldots, \theta_N$ are assumed to be *distinct* members of Θ. The c_i and θ_i, $i = 1, \ldots, N$ will be called respectively the *coeffients* and *support points* of the finite mixture.

According to [6] we have the identifiability criteria for finite mixtures. The following formal definition states that the class of finite mixtures $\widetilde{\boldsymbol{\mathcal{H}}}$ is identifiable if and only if all members of $\widetilde{\boldsymbol{\mathcal{H}}}$ are distinct.

Definition 4.3. Let $\widetilde{\boldsymbol{\mathcal{H}}}$ be the class of finite mixtures on $\boldsymbol{\mathcal{F}}$. $\widetilde{\boldsymbol{\mathcal{H}}}$ is ***identifiable*** if and only if

$$\sum_{i=1}^{N} c_i F(\cdot,\theta_i) = \sum_{i=1}^{\widehat{N}} \widehat{c}_i F(\cdot,\widehat{\theta}_i)$$

implies $N = \widehat{N}$ and for each i, $1 \le i \le N$, there is j, $1 \le j \le N$, such that $c_i = \widehat{c}_j$ and $\theta_i = \widehat{\theta}_j$.

Lemma 4.4 (Setiawaty [5]). *Let $\widetilde{\mathcal{H}}$ be the class of finite mixtures on $\mathcal{F}$. $\widetilde{\mathcal{H}}$ is identifiable if and only if*

$$\sum_{i=1}^{N} c_i F(\cdot, \theta_i) = \sum_{i=1}^{\widehat{N}} \widehat{c}_i F(\cdot, \widehat{\theta}_i) \qquad \Longrightarrow \qquad N = \widehat{N}, \quad \sum_{i=1}^{N} c_i \delta_{\theta_i} = \sum_{i=1}^{N} \widehat{c}_i \delta_{\widehat{\theta}_i},$$

where δ_θ denotes the Dirac distribution of a point mass at θ.

5. Identifiability of Hidden Markov Models

Let $\{(X_t, Y_t)\}$ be a hidden Markov model with representation $\phi = (K, A, \pi, \theta) \in \Phi_K$. From section 2, the parameters A, π and θ satisfy :

$$A = (\alpha_{ij}), \quad \alpha_{ij} \geq 0, \quad \sum_{j=1}^{K} \alpha_{ij} = 1, \quad i, j = 1, \dots, K$$
$$\pi = (\pi_i), \quad \pi_i \geq 0, \quad i = 1, \dots, K, \quad \sum_{i=1}^{K} \pi_i = 1$$
$$\theta = (\theta_i)^T, \quad \theta_i \in \Theta, \quad i = 1, \dots, K.$$

Notice that $\theta_1, \theta_2, \dots, \theta_K$ *need not all to be distinct.*

Under ϕ, for any $n \in \boldsymbol{N}$, the joint density function of $Y_1, \dots, Y_n$ is

$$p_\phi(y_1, \dots, y_n) = \sum_{x_1=1}^{K} \cdots \sum_{x_n=1}^{K} \pi_{x_1} f(y_1, \theta_{x_1}) \prod_{t=2}^{n} \alpha_{x_{t-1}, x_t} f(y_t, \theta_{x_t}). \tag{5.1}$$

Let

$$Q_\phi = \sum_{x_1=1}^{K} \cdots \sum_{x_n=1}^{K} \pi_{x_1} \prod_{t=2}^{n} \alpha_{x_{t-1}, x_t} \delta_{(\theta_{x_1}, \dots, \theta_{x_n})}, \tag{5.2}$$

then (5.1) can be written as

$$p_\phi(y_1, y_2, \dots, y_n) = \int_{\Theta^n} f(y_1, \zeta_1) \cdots f(y_n, \zeta_n) Q_\phi(d\zeta_1, \dots, d\zeta_n). \tag{5.3}$$

Equations (5.1), (5.2) and (5.3) assert that, for $n = 1$, p_ϕ is a finite mixture with non-negative coefficients $\pi_1, \dots, \pi_K$ and may not be distinct support points $\theta_1, \dots, \theta_K$. For $n \geq 2$, p_ϕ is a finite mixture of product measures with non-negative coefficients $\left(\pi_{x_1} \prod_{t=2}^{n} \alpha_{x_{t-1}, x_t}\right)$ and may not be distinct support points $(\theta_{x_1}, \dots, \theta_{x_n})$, for $x_1, \dots, x_n \in \{1, \dots, K\}$.

In order to apply the identifiability of finite mixtures to hidden Markov models, Definition 4.3 has to be relaxed to allow the above possibilities.

Definition 5.1. Let $\boldsymbol{\mathcal{F}} = \{F(\cdot, \theta) : \theta \in \Theta\}$ be a family of one dimensional distribution functions defined on $\mathcal{Y}$ indexed by $\theta \in \Theta$. Let

$$\widehat{\boldsymbol{\mathcal{H}}} = \left\{ H(\cdot) \ : \ H(\cdot) = \sum_{i=1}^{K} c_i F(\cdot, \theta_i), \right.$$
$$\left. c_i \geq 0, \ \theta_i \in \Theta, \ i = 1, 2, \dots, K, \ \sum_{i=1}^{K} c_i = 1, K \in \boldsymbol{N} \right\} \tag{5.4}$$

$\widehat{\boldsymbol{\mathcal{H}}}$ is ***identifiable*** if and only if

$$\sum_{i=1}^{K} c_i F(\cdot, \theta_i) = \sum_{i=1}^{\widehat{K}} \widehat{c}_i F(\cdot, \widehat{\theta}_i) \quad \Longrightarrow \quad \sum_{i=1}^{K} c_i \delta_{\theta_i} = \sum_{i=1}^{\widehat{K}} \widehat{c}_i \delta_{\widehat{\theta}_i}. \tag{5.5}$$

where δ_θ denotes the Dirac distribution of a point mass at θ.

Remarks 5.2. In every expression of

$$H(\cdot) = \sum_{i=1}^{K} c_i F(\cdot, \theta_i) \in \widehat{\boldsymbol{\mathcal{H}}},$$

the parameters $\theta_1, \dots, \theta_K$ *need not all to be distinct.*

Next lemma shows the relation between Definition 4.3 and Definition 5.1.

Lemma 5.3. *$\widehat{\boldsymbol{\mathcal{H}}}$ is identifiable according to Definition 5.1 if and only if $\widetilde{\boldsymbol{\mathcal{H}}}$ is identifiable according to Definition 4.3.*

Proof :
Necessity :
Assume that $\widehat{\boldsymbol{\mathcal{H}}}$ is identifiable according to Definition 5.1. We will prove that $\widetilde{\boldsymbol{\mathcal{H}}}$ is identifiable according to Definition 4.3. Suppose

$$\sum_{i=1}^{K} c_i F(\cdot, \theta_i) = \sum_{i=1}^{\widehat{K}} \widehat{c}_i F(\cdot, \widehat{\theta}_i), \tag{5.6}$$

where :

$$c_i > 0, \quad i = 1, \dots, K, \quad \sum_{i=1}^{K} c_i = 1$$
$$\widehat{c}_i > 0, \quad i = 1, \dots, \widehat{K} \quad \sum_{i=1}^{\widehat{K}} \widehat{c}_i = 1$$
θ_i are distinct for $i = 1, \dots, K$
$\widehat{\theta}_i$ are distinct for $i = 1, \dots, \widehat{K}$.

By Definition 5.1, equation (5.6) implies

$$\sum_{i=1}^{K} c_i \delta_{\theta_i} = \sum_{i=1}^{\widehat{K}} \widehat{c}_i \delta_{\widehat{\theta}_i}. \tag{5.7}$$

Since $c_i > 0$ and θ_i are distinct for $i = 1, \dots, K$, then according to [5], $\widehat{K} \geq K$. On the otherhand, since $\widehat{c}_i > 0$ and $\widehat{\theta}_i$ are distinct for $i = 1, \dots, \widehat{K}$, then according to [5], we also have $K \geq \widehat{K}$. Hence, we have $K = \widehat{K}$ and by (5.7),

$$\sum_{i=1}^{K} c_i \delta_{\theta_i} = \sum_{i=1}^{K} \widehat{c}_i \delta_{\widehat{\theta}_i}.$$

By Lemma 4.4, $\widetilde{\boldsymbol{\mathcal{H}}}$ is identifiable according to Definition 4.3.

Sufficiency :
Assume that $\widetilde{\boldsymbol{\mathcal{H}}}$ is identifiable according to Definition 4.3. We will prove that $\widehat{\boldsymbol{\mathcal{H}}}$ is identifiable according to Definition 5.1. Suppose

$$\sum_{i=1}^{K} c_i F(\cdot, \theta_i) = \sum_{i=1}^{\widehat{K}} \widehat{c}_i F(\cdot, \widehat{\theta}_i), \tag{5.8}$$

where :

$$\begin{array}{l}
c_i \geq 0, \quad i = 1, \dots, K, \quad \displaystyle\sum_{i=1}^{K} c_i = 1 \\
\widehat{c}_i \geq 0, \quad i = 1, \dots, \widehat{K}, \quad \displaystyle\sum_{i=1}^{\widehat{K}} \widehat{c}_i = 1 \\
\theta_i \text{ need not all to be distinct, for } i = 1, 2, \dots, K \\
\widehat{\theta}_i \text{ need not all to be distinct, for } i = 1, \dots, \widehat{K}.
\end{array}$$

Let

$$\begin{aligned}
F_+ &= \{i : c_i > 0, i = 1, \dots, K\} \\
\widehat{F}_+ &= \{i : \widehat{c}_i > 0, i = 1, \dots, \widehat{K}.
\end{aligned}$$

Let r be the number of distinct θ_i, $i \in F_+$ and $\widehat{r}$ be the number of distinct $\widehat{\theta}_i$, $i \in \widehat{F}_+$. Without loss of generality, suppose that $\theta_1, \dots, \theta_r$ are distinct and also $\widehat{\theta}_1, \dots, \widehat{\theta}_{\widehat{r}}$. Let

$$\begin{aligned}
R_i &= \{j : j \in F_+, \theta_j = \theta_i\}, \quad i = 1, \dots, r \\
\widehat{R}_i &= \{j : j \in \widehat{F}_+, \widehat{\theta}_j = \widehat{\theta}_i\}, \quad i = 1, \dots, \widehat{r}.
\end{aligned}$$

Equation (5.8) then can be written as

$$\sum_{i=1}^{r} a_i F(\cdot, \theta_i) = \sum_{i=1}^{\widehat{r}} \widehat{a}_i F(\cdot, \widehat{\theta}_i), \tag{5.9}$$

where

$$a_i = \sum_{j \in R_i} c_j \qquad \text{and} \qquad \widehat{a}_i = \sum_{j \in \widehat{R}_i} \widehat{c}_j.$$

Since $a_i > 0$ and θ_i are distinct for $i = 1, \ldots, r$; and $\widehat{a}_i > 0$ and $\widehat{\theta}_i$ are distinct for $i = 1, \ldots, \widehat{r}$, then by Definition 4.3, equation (5.9) implies $r = \widehat{r}$ and

$$\sum_{i=1}^{r} a_i \delta_{\theta_i} = \sum_{i=1}^{r} \widehat{a}_i \delta_{\widehat{\theta}_i}. \tag{5.10}$$

But this is equivalent with

$$\begin{aligned} \sum_{i=1}^{r} \sum_{j \in R_i} c_j \delta_{\theta_j} &= \sum_{i=1}^{\widehat{r}} \sum_{j \in \widehat{R}_i} \widehat{c}_j \delta_{\widehat{\theta}_j} \\ \sum_{i \in F_+} c_i \delta_{\theta_i} &= \sum_{i \in \widehat{F}_+} \widehat{c}_i \delta_{\widehat{\theta}_i}. \end{aligned} \tag{5.11}$$

Since $c_i = 0$, for $i \notin F_+$ and also $\widehat{c}_i = 0$, for $i \notin \widehat{F}_+$, then by (5.11)

$$\sum_{i=1}^{K} c_i \delta_{\theta_i} = \sum_{i=1}^{\widehat{K}} \widehat{c}_i \delta_{\widehat{\theta}_i}.$$

Hence, $\widehat{\boldsymbol{\mathcal{H}}}$ is identifiable according to Definition 5.1. ■

Remarks 5.4. As a consequence of Lemma 5.3, all the results of identifiability in section 2.1 are now applicable for hidden Markov models. So from now on, when we say $\widehat{\boldsymbol{\mathcal{H}}}$ is identifiable, we mean it in the sense of Definition 5.1.

References

[1] Everitt B.S and Hand D.J. 1981. *Finite Mixture Distributions*. Chapman and Hall, London.

[2] Ryden, T. 1996. On recursive estimation for hidden Markov models. *Stochastic Processes and their Applications*, 66, 79-96.

[3] Setiawaty, B. 2002. Equivalent representations of hidden Markov models. *Journal of Mathematics and its Aplications*, Volume 2, No.1, 1-12.

[4] Setiawaty, B. 2003. Characteristics of a true parameter of a hidden Marcov model. *Journal of Mathematics and its Applications*, Volume 2, no. 2: 43-48.

[5] Setiawaty, B. 2004. A survey on the identifiability of finite mixtures. *Journal of Mathematics and its Applications*, Volume 3, no. 2.

[6] Teicher, H. 1963. Identifiability of finite mixtures. *Ann. Math. Statist.*, 34: 1265-1269.